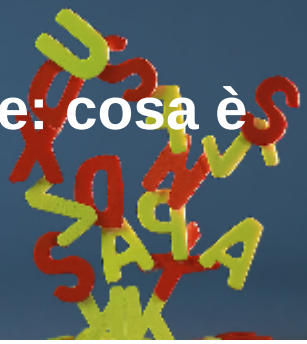
Digeat N.7 - 19 Settembre 2025

La qualità del dato personale: cosa è davvero imperativo?

Di Giovanni Ferorelli



Abstract

La qualità dei dati personali è un tema centrale, capace di incidere direttamente su decisioni pubbliche, servizi e diritti individuali. Nel presente articolo, dopo una breve riflessione su cosa debba intendersi con “qualità” del dato personale, si riflette sugli elementi fondanti la qualità del dato e su quale approccio l’organizzazione dovrebbe avere per garantire tale requisito. Non basta affidarsi alla tecnologia, ma occorrono competenze, progettazione e processi efficienti, nonché responsabilità chiare. L’obiettivo non è la perfezione assoluta, bensì poter contare su dati di un livello di qualità adeguato in base al rischio e al contesto.

Indice

- Introduzione
- La “qualità” dei dati personali
- Come garantire la qualità del dato
- L’esattezza del dato personale quale concetto per certi versi relativo

Introduzione

La qualità dei dati, oggi più che mai, gioca un ruolo cruciale: i dati orientano le azioni governative, sull’analisi dei dati si basano scelte e strategie aziendali, i dati vengono trattati per fornire servizi di varia natura, ecc. Specialmente nella odierna società digitalizzata, è facile intuire quanto la qualità del dato, nel bene o nel male, possa avere riflessi sulla vita di ciascuno di noi e sulle future generazioni.

Si pensi alla qualità dei dati meteorologici, giusto per fare un esempio. Se imprecisi, possono compromettere le previsioni del tempo, mettendo a rischio la sicurezza di cittadini e lavoratori, oltre che il buon andamento di attività, programmi e processi aziendali. E si immagini ora, restando in tema, che dati imprecisi vengano utilizzati per sviluppare un modello di intelligenza artificiale predittivo e per fare previsioni meteorologiche a lungo termine. Verrebbe compromessa la qualità – e dunque l’utilità – dello stesso modello e dei relativi output, con potenziali gravi e perduranti conseguenze a catena, a danno di varie categorie di soggetti, qualora sul modello e sui relativi output si dovesse fare affidamento[1].

> L’argomento è estremamente vasto e pone così tanti quesiti di natura giuridica ed etica che anche solo cercare di elencarli sarebbe in questa sede un tentativo velleitario. Perciò, di seguito mi limiterò a fare qualche riflessione sulla qualità del dato personale, cercando di offrire spunti sull’approccio consapevole e strutturato da adottare in relazione alla qualità del dato.

Per far ciò, però, è innanzitutto necessario stabilire cosa dobbiamo intendere, almeno nel presente articolo, con “qualità” del dato.

La “qualità” dei dati personali

Ebbene, in ambito data protection, il concetto di qualità del dato è generalmente accostato, anche da parte del Garante per la protezione dei dati personali[2], all’esattezza dei dati personali – dunque al principio di cui all’art. 5, lett. d) del GDPR – alla loro accuratezza e affidabilità.

Il citato art. 5 non si limita a prevedere che i dati personali debbano essere “**esatti e, se necessario, aggiornati**”, ma specifica anche che “devono essere adottate tutte le misure **ragionevoli** per cancellare o rettificare tempestivamente i dati inesatti **rispetto alle finalità per le quali sono trattati**”. Un requisito, dunque, l’esattezza del dato, che è strettamente correlato alle finalità perseguite, ed è anche in base a queste che, pertanto, occorre individuare quali misure debbano ritenersi ragionevoli in un determinato contesto. Perciò, ogni valutazione circa l’esattezza del dato andrebbe fatta sempre avendo bene a mente quelle che sono le finalità che si intende perseguire tramite il trattamento di un determinato dato.

E allora, nell’ambito di cui si tratta, usare i termini “qualità” ed “esattezza” quali sinonimi può essere riduttivo, e non far cogliere pienamente le altre caratteristiche che un dato (o un set di dati) di qualità deve presentare.

> Ecco, dunque, che la definizione “data quality”, offerta dallo standard internazionale ISO 8000-2:2022, può tornare utile ai nostri fini, in quanto coerente con i principi di cui sopra. Ai sensi di tale norma, è possibile intendere la qualità dei dati come il **grado in cui un insieme di caratteristiche intrinseche dei dati soddisfa le esigenze o le aspettative espresse, generalmente implicite o obbligatorie**[3].

Ai fini del presente articolo, si intenda dunque la qualità del dato come un concetto che richiede la sussistenza di almeno due ordini di requisiti, ossia:

1. idoneità a soddisfare le finalità perseguite;
2. esattezza, accuratezza, completezza,

dove, a ben vedere, l’esattezza del dato può essere intesa come caratteristica fondamentale – ma non unica[4] – affinché il dato sia idoneo a soddisfare determinate finalità: è evidente come il perseguimento di queste possa essere pregiudicato da dati inesatti.

Ma come garantire la qualità del dato? Nel paragrafo successivo provo a fornire qualche spunto.

Come garantire la qualità del dato

Per garantire la qualità del dato personale, nell’accezione di cui sopra, occorre pensare al dato nell’arco del suo intero ciclo di vita, anzi, ancor prima della sua generazione/acquisizione. A titolo esemplificativo:

- in fase di progettazione di un nuovo servizio, ci si deve chiedere quali dati occorre generare o acquisire, e con quali modalità e secondo quali processi, per poter perseguire determinate finalità (e questo riguarda tutte le categorie di dati personali, anche quelli che possono apparire intrinsecamente inesatti come, ad esempio, le opinioni personali);

- una volta acquisiti, occorre organizzare, strutturare e archiviare i dati secondo logiche e misure che ne garantiscano la qualità nel tempo (dunque anche la sicurezza);
- durante tutto il ciclo di vita dei dati, l'organizzazione deve adottare procedure finalizzate a garantire l'aggiornamento e la rettifica degli stessi e se, come spesso succede, nel trattamento sono coinvolte più organizzazioni, le responsabilità a riguardo devono essere ben definite e correttamente distribuite. Ciò vuol dire anche che l'organizzazione, e tutte le persone coinvolte nel trattamento, devono essere in grado di riconoscere e gestire correttamente e ai sensi di legge eventuali richieste di rettifica da parte degli interessati. Vale la pena ricordare che il diritto di rettifica ha, nella società odierna, una importanza fondamentale: non è solo garantito dal GDPR, ma è anche richiamato dalla Carta dei diritti fondamentali dell'Unione Europea che, all'art. 8, par. 2, sancisce che "**ogni individuo ha il diritto di accedere ai dati raccolti che lo riguardano e di ottenerne la rettifica**";
- i dati devono essere utilizzati in modo tale da non pregiudicarne la qualità.

Un'organizzazione può disporre di dati con un indice di accuratezza del 100%, ma se non sono archiviati in modo corretto, quale vantaggio se ne trae? O, al contrario, un'organizzazione dispone delle migliori tecnologie per poter analizzare in maniera efficiente grandi moli di dati, ma se questi dati sono ridondanti, oppure, per qualsiasi motivo imprecisi, a cosa serve la potenza di calcolo o, comunque, la tecnologia super avanzata?

> È per questi motivi che la qualità del dato non può che essere pensata come il risultato dell'interazione tra loro di vari fattori, tra cui: valutazioni e scelte, sia durante il trattamento sia in fase preliminare, su molteplici aspetti; operazioni; procedure; strumenti. In definitiva, possiamo affermare che la qualità del dato dipende dall' **interazione tra persone e tecnologia** e che, in diversa misura, **a partire dai vertici apicali aziendali** – ma il discorso potrebbe essere esteso anche al di fuori del mondo imprenditoriale – **tutti, all'interno dell'organizzazione, contribuiscono a garantire la qualità del dato, e ne sono in qualche modo responsabili**.

In termini più operativi, di seguito riporto alcune preziose indicazioni rese dall'European Data Protection Board[5] a proposito dell'esattezza dei dati personali quale elemento fondamentale nella realizzazione di trattamenti che rispettino i principi della protezione dei dati personali fin dalla progettazione e per impostazione predefinita (data protection by design and by default):

- fonte dei dati – le fonti dei dati personali dovrebbero essere affidabili in termini di esattezza dei dati;
- grado di esattezza – ciascun elemento dei dati personali deve essere il più esatto possibile in base alle necessità delle finalità specifiche;
- esattezza misurabile – occorre ridurre il numero di falsi positivi/negativi, per esempio le distorsioni generate nell'ambito delle decisioni automatizzate e dell'intelligenza artificiale;
- verifica – a seconda della natura dei dati, e in relazione alla frequenza delle relative modifiche, il titolare dovrebbe verificare la correttezza dei dati personali presso l'interessato prima del trattamento e nelle sue diverse fasi (per esempio rispetto ai requisiti di età);
- cancellazione/rettifica – il titolare dovrebbe cancellare o rettificare tempestivamente i dati inesatti e, in particolare, agevolare questa procedura se gli interessati sono o erano minori e successivamente desiderano eliminare i suddetti dati personali;
- evitare la propagazione di errori – i titolari dovrebbero attenuare l'effetto di un errore accumulato nella catena di trattamento;
- accesso – gli interessati dovrebbero ricevere informazioni sui dati personali e disporre di un accesso efficace agli stessi, ai sensi degli articoli da 12 a 15 del RGPD, per controllarne l'esattezza e apportare le rettifiche ove necessario;
- esattezza permanente – i dati personali dovrebbero essere esatti in tutte le fasi del trattamento e nelle fasi critiche dovrebbero essere effettuate verifiche di esattezza;

- aggiornamento – i dati personali sono aggiornati qualora ciò sia necessario per la specifica finalità;
- progettazione dei dati – impiego di caratteristiche organizzative e tecnologiche di progettazione per ridurre le eventuali inesattezze, per esempio proponendo scelte concise e predeterminate anziché campi a testo libero.

Tali indicazioni non esauriscono gli elementi da valutare in relazione all'esattezza dei dati, e ragionamenti e misure specifici andrebbero fatti e adottati caso per caso.

> Ma una cosa è certa: **è di fondamentale importanza, e di natura propedeutica, l'individuazione del grado di esattezza che si vuole pretendere in relazione a determinati dati o set di dati**.

Detto altrimenti, quale percentuale di errore nell'esattezza dei dati si è disposti a (o meglio, è ragionevole) accettare?

L'esattezza del dato personale quale concetto per certi versi relativo

Per cercare una risposta a questa domanda, partiamo da alcuni possibili scenari, tenendo ora in considerazione quali possano essere i rischi derivanti da trattamenti di dati inesatti/di scarsa qualità per gli interessati (il GDPR richiede un approccio basato sul rischio per i diritti e le libertà degli interessati).

- **Es. 1**: l'indirizzo e-mail di Rocco al quale inviare la newsletter mensile sulle novità della Cava dei Dinosauri di Altamura non è corretto, e Rocco per mesi non riceve alcuna comunicazione al riguardo.
- **Es. 2**: l'anagrafica di un cliente non è aggiornata, e i primi tentativi di recapito nel computer vanno a vuoto, ritardando di un mese la consegna.
- **Es. 3**: a causa di un difetto nella procedura di raccolta dei dati mediante un questionario anamnestico, viene erroneamente indicata la sussistenza di una grave malattia pregressa di cui Tizio sarebbe stato affetto, con la conseguente esclusione dalla copertura assicurativa richiesta.
- **Es. 4**: esausto, dopo 48 ore consecutive di servizio a causa di assenza di personale, un infermiere scambia inavvertitamente le etichette applicate sulle provette di sangue. Di conseguenza, gli esiti degli esami vengono attribuiti al paziente sbagliato, determinando una diagnosi patologica errata in un soggetto sano, mentre al paziente realmente affetto dalla patologia non viene diagnosticata la malattia in tempo utile per un trattamento curativo.

Dagli esempi di cui sopra – che possono prendere le forme e le sfumature più diverse a seconda delle circostanze, emerge chiaramente come l'inesattezza dei dati personali, a seconda dei contesti in cui avviene il trattamento, può arrecare alle persone danni di varia natura e gravità, che possono andare dal semplice fastidio per non aver ricevuto una newsletter fino alla morte di un paziente per errata diagnosi. E pretendere di dare all'indirizzo e-mail per l'invio delle newsletter e alle etichette degli esami del sangue la medesima attenzione è evidentemente irragionevole.

> Ne consegue un principio guida essenziale: **è il rischio per i diritti e le libertà degli interessati derivante da un trattamento di dati inesatti che dovrebbe orientare l'organizzazione – con il supporto del DPO[6] – nell'individuazione del grado di esattezza che si vuole pretendere in relazione a determinati dati o set di dati** e, di conseguenza, nell'individuazione delle misure tecniche e organizzative (compresi i controlli) adeguate, per assicurare che quel grado di esattezza venga rispettato.

Ragionevolezza, quindi, tornando al principio di esattezza di cui all'art. 5 del GDPR, in relazione alle finalità e al rischio.

In conclusione, e cercando di rispondere alla domanda introduttiva, ciò che è davvero imperativo è la necessità di costruire processi e formare persone in modo tale da garantire che i dati presentino un grado di qualità ragionevolmente adeguato in base al contesto di riferimento, alle finalità e ai rischi per i diritti e le libertà degli interessati. Non abbiamo bisogno di dati "perfetti", ma di processi che generino dati di qualità "adeguata".

NOTE

[1] La qualità del dato è un requisito fondamentale anche nell'ambito dell'IA, e sotto diversi aspetti. Fermo restando che il Regolamento (UE) 2024/1689 (Regolamento sull'intelligenza artificiale, o Artificial Intelligence Act – AI Act) lascia impregiudicato il GDPR e, quindi, il principio di esattezza dei dati va rispettato – per certi versi ciò presenta una sfida – anche con riferimento ai dati trattati mediante sistemi di IA, qui ci si limita a evidenziare che il concetto di "data quality" ricorre molto frequentemente nell'AI Act. Inoltre, giova sottolineare che tra i sette principi etici enunciati negli Orientamenti etici per un'IA affidabile del 2019, elaborati dall'AI HLEG indipendente nominato dalla Commissione – e richiamati nell'AI Act al considerando 27 – figura il principio della "**vita privata e governance dei dati**", con cui si intende che **i sistemi di IA sono sviluppati e utilizzati nel rispetto delle norme in materia di vita privata e protezione dei dati, elaborando al contempo dati che soddisfino livelli elevati in termini di qualità e integrità**.

[2] Si vedano, ad esempio, l'intervento della Vice Presidente del Garante per la protezione dei dati personali, Prof.ssa Ginevra Cerrina Feroni, a proposito di "Fisco e privacy, alla ricerca dell'equilibrio tra diritti e lotta all'evasione" pubblicato il 22 maggio 2025 su AgendaDigitale e disponibile anche a questo link, e la Relazione del Garante per la protezione dei dati personali sull'attività 2024, pagg. 101, 213.

[3] Più precisamente, lo standard internazionale ISO 8000-2:2022 – documento che definisce termini, relativi alla qualità dei dati, usati in parti della serie ISO 8000 – definisce così "**Data quality**": "degree to which a set of inherent characteristics of data fulfils requirements", dove con il termine "requirements" deve intendersi: "need or expectation that is stated, generally implied or obligatory".

[4] A tal proposito, è interessante richiamare l'art. 6, comma 1, del decreto legislativo 4 marzo 2013, n. 33, che, in relazione al requisito della "qualità delle informazioni" richiama i concetti di **integrità**, **aggiornamento**, **completezza**, **tempestività**, semplicità di **consultazione**, **comprensibilità**, **omogeneità**, facile **accessibilità** e **conformità** ai documenti originali in possesso dell'amministrazione, nonché **indicazione della loro provenienza** e **riutilizzabilità**. Più precisamente, il citato comma prevede che: "Le pubbliche amministrazioni garantiscono la qualità delle informazioni riportate nei siti istituzionali nel rispetto degli obblighi di pubblicazione previsti dalla legge, assicurandone l'integrità, il costante aggiornamento, la completezza, la tempestività, la semplicità di consultazione, la comprensibilità, l'omogeneità, la facile accessibilità, nonché la conformità ai documenti originali in possesso dell'amministrazione, l'indicazione della loro provenienza e la riutilizzabilità secondo quanto previsto dall'articolo 7".

[5] EDPB, "Linee guida 4/2019 sull'articolo 25 Protezione dei dati fin dalla progettazione e per impostazione predefinita", consultabili qui.

[6] Il DPO, infatti, ai sensi dell'art. 39, par. 2 del GDPR, è tenuto a eseguire i propri compiti considerando debitamente i rischi inerenti al trattamento.